7

To: "Philip Inghelbrecht" <inghelbrecht@google.com>
From: "Jeremy Doig" <jeremydo@google.com>
Cc: "salar@google.com" <salar@google.com>, "Peter Chane" <peterch@google.com>, "Jennifer Feikin" <jfeikin@google.com>, "Alex Ellerson" <ellerson@google.com>, "David Eun" <deun@google.com>
Bcc:
Received Date: 2006-11-22 19:38:02 GMT
Subject: Re: Audible Magic: perhaps not ideal for Google Video / YouTube?

I talked to Franck at YT about this yesterday - notwithstanding the technical shortcomings of AM, the business model generally worries me - I think this is core technology we need to build and control ourselves. So the current approach is to only use AM as little as possible and proceed with building our own database at the same time. This would require us either getting all the source files (preferred) or distributing fingerprinting tools which content providers use. Our inhouse audio fingerprinting currently is at 6 second intervals and the video is 1-2 minutes, but we're working on improving that significantly (it hasn't been a focus area until recently).

(And I always liked shazam as a product)

On 11/22/06, Philip Inghelbrecht <inghelbrecht@google.com> wrote:
>
> Jeremy, Salar, Peter, Jennifer, Alex and Dave:
> (Sorry if this mail is somewhat long but pls give it a read when you have
> 3min over the wkd)
>
> I understand that YouTube intends to deploy Audible Magic (AM) as a
> copyright detection tool. Having founded a similar company a few years ago
> (Shazam Entertainment Ltd - www.shazam.com), I know and understand these
> fingerprint technologies extremely well. AM is a second-tier provider
> and adopting YouTube's vendor could lock Google into an inferior algorithm
> and/or business model which may harm us in the longer term. I have
> shortlisted some reasons below.
>
> - Accuracy: AM will need at least 20-30sec with a possible <65%
> accuracy (and likely high false positive rate). In contrast, the Shazam
> technology was built to recognize 10sec of music that had undergone cellular
> (e.g. GSM) compression. When the technology is used with audio of
> much higher quality (et as low as 8kHz WAV), it will attain an accuracy of
> 99+% using only a few seconds of source material.
> - Scalability: a good fingerprinting technology should be able to
> handle hundreds of queries per second. AM is in the low teens.
> ClearChannel monitors its 1,500+ US radio stations 24x7 using Shazam with
> spare capacity (and it is a well-known fact that the Shazam technology can
> be tweaked to operate even faster).
> - Robustness: most technologies will fail here. E.g. if users were
> to speed up music with 0.5% (=easy), it will be inaudible to the
> human ear but will slip through the AM detection system. However, Shazam was
> built to deal with speed variation (with or without pitch control applied!)
> and would not suffer from these distortions.
> - Database size : most technologies fall over when the matching
> database exceeds 100,000 tracks (and AM notoriously does so). Shazam
> currently operates with 3+MM songs and just like Google, there is no
> scalability issue on the horizon.
> - Matching database: AM asks the labels to submit the reference

> songs. Companies like Shazam and Gracenote have this readily available and 7-0002
> spend millions of dollars each year to keep it up-to-date. Furthermore,
> since Shazam operates in 20+ countries worldwide, its music selection will
> be unparallelled and be fully aligned with the international expansion of
> GV/YT.
> - Granularity: in addition to only needing a few seconds of music,
> Shazam can also pinpoint the exact section of the song (to the nearest
> millisecond). This will be important for Google in the future to correctly
> detect & assess music use in mash-up videos.
> - Business model: AM and Gracenote require an ASP model. Shazam
> doesn't lock its partner into such model but let them run the software
> inhouse. Large customers get full access to the binary code and hardware
> setup (clustered nodes, a very similar approach to Google).
>
> There are five main players in the audio fingerprinting market: Gracenote(who acquired the technology from Philips),
> AM (who purchased the technology from Musclefish), BDS (developed in the
> 70s and aging out), Mediaguide (developed by Connexus) and Shazam(developed in-house). Snocap is also often mentioned but they simply
> license from Gracenote/Phillips.
>
> Over the last 3-4 years, AM was the only company to actively focus on
> copyright protection tools (its weaker technology could not compete in the
> more lucrative markets) and has as such developed an incredible amount of
> goodwill within the music industry. AM has also gained market traction by
> licensing their technology extremely cheaply (or: quality comes at a cost).
> In light of the details bullet-listed above, that doesn't make AM the best
> vendor!
>
> Last but not least: the Shazam technology was snapped up by BMI (the
> world's largest performance rights society -
> http://www.bmi.com/news/200508/20050830a.asp) and as such I can consider
> myself a neutral party to this strong recommendation: we should review all
> five technologies above and (logically?) switch to BMI (Shazam) or perhaps
> even Gracenote (which I consider a strong #2).
>
> Thanks for hearing me out & feel free to forward to the relevant YouTube
> folks. [p]
>
> PS - Google currently employs 6 ex-Shazam employees, my co-founder Chris
> Barton included. Furthermore, Anil Hansjee (who joined the corp dev team in
> the UK) was an observer on the board. I strongly suspect that all 7 people
> would confirm the above (and none have a vested interest to do so).
>
>
> --
> Philip Inghelbrecht
> Strategic Partner Development
> Google Inc
> [redacted]
